

Konspekt prezentacji

1. Temat pracy dyplomowej, charakterystyka danych
2. Czym jest przebieg czasowy, przykłady
3. Cele
 - a. Stworzenie własnego klasyfikatora
 - b. Porównanie stworzonego klasyfikatora z klasyfikatorem przeznaczonym do analizy przebiegów czasowych (miarodajny punkt odniesienia)
4. Możliwe podejścia do analizy przebiegów czasowych
 - a. Klasyczna regresja
 - b. „Wygładzanie” (Smoothing)
 - c. Modele ARIMA (Autoregressive Moving Average Models)
 - d. „Analiza spektrum” <spektralna?>(Spectral Analysis)
 - e. State-Space Models
5. SAX – co to jest, przykłady
6. Jak wrzucić SAXa do problemu jaskry – i dalej – do klasyfikowania podobnych problemów dotyczących przebiegów czasowych

Ad. 4a)

Podejście najbardziej bazowe, prosta koncepcja. Wyjściem metody jest funkcja liniowa, która w dobry sposób estymuje funkcje monotoniczne, a przy sinusoidalnych sprawia wiele kłopotu. Realny cel, do którego mogłaby posłużyć w tym przypadku jest wyznaczenie jakiegoś trendu, co jest zdecydowanie niewystarczające, ponieważ potrzebna jest znacznie większa dokładność

Ad. 4b)

Podobnie jak regresja, metoda ta skupia się na specyficznym atrybucie, najczęściej w kontekście długoterminowym, zatem do przedmiotu tej pracy nie jest najlepszym wyborem, ponieważ posiada niewystarczającą ilość informacji o dynamice badanego szeregu czasowego

Ad. 4c)

Modele ARIMA wnoszą do analizy danych czasowych nową wartość – dotychczasowe podejścia nie brały pod uwagę tego, że wartości historyczne mogą wpływać na wartości uzyskane w późniejszym okresie (na tym samym lub nawet innych atrybutach). Niekonwencjonalny jak dotąd wzór ARIMA bazuje na idei przedstawienia obecnej wartości funkcji na podstawie p poprzednich wartości. Przykładowa funkcja wygląda następująco:

$$x_t = x_{t-1} - 1.4x_{t-2} + \omega_t$$

Jednak modele ARIMA nie są koncepcyjnie stworzone w celu klasyfikowania danych czasowych, czy podejmowania na ich podstawie jakichś decyzji. Służą one raczej do prognozowania wartości przyszłych dla zadanego szeregu wartości, szczególnie w odniesieniu do krótkich przebiegów czasowych.

Ad. 4d)

Analiza spektrum zwraca uwagę na kolejną charakterystyczną cechę pewnej rodziny przebiegów czasowych – a mianowicie na powtarzalność lub mówiąc inaczej – okresowość. W tej metodzie dominuje spojrzenie na wariacje funkcji sinus oraz cosinus – produkując ich liniowe kombinacje w celu jak najdoskonalszego przybliżenia do bieżących danych. Ze względu na liczne modyfikacje tej metody (zastosowanie Dyskretnej Transformaty Fourier’a, uwzględnienie gęstości spektrum, filtrów liniowych) oraz konfigurowalne atrybuty (amplituda, częstotliwość) można sądzić, że ta metoda sprawdziłaby się jako model decyzyjny dla przedmiotu tej pracy, jednak w mojej opinii jest metodą dość skomplikowaną matematycznie i prawdopodobnie będzie wymagająca jeśli chodzi o implementację.

Ad. 4e)

<Czy na pewno potrzebna nam 5. metoda? Zdaje się być najtrudniejsza – zarówno jeśli chodzi o zrozumienie jak i wdrożenie. Za drugim podejściem wciąż jej nie rozumiem>